Ripplet II Transform and Higher Order Cumulants from R-fMRI data for Diagnosis of Autism

Shadi Sartipi, Hashem Kalbkhani, and Mahrokh G. Shayesteh

Department of Electrical Engineering, Urmia University, Urmia, Iran

Emails: {st_sh.sartipi, h.kalbkhani, m.shayesteh}@urmia.ac.ir

Abstract

Diagnose of autism spectral disorder (ASD) as a mental disorder by machine learning algorithms has attracted many attentions. Finding biomarkers from the rest state functional magnetic resonance imaging (R-fMRI) data is one of the common methods used for classifying ASD and normal healthy person (HP). This paper presents Eickhoff-Zilles (EZ) atlas to evaluate time courses for 20 ASDs and 16 HPs in 116 regions of interest (ROIs). To extract the effective features for classification, Ripplet II transform and higher order cumulants are proposed. Then, two sample t-test is employed to select the discriminative features for classification. After normalizing the selected feature vector, the data are classified by support vector machine (SVM). The results show that the proposed method achieves 91.67% accuracy which outperforms previous works.

1. Introduction

Autism spectral disorder (ASD) is a mental disability which causes disturbance in social skills, speech, and verbal communications [1]. The signs of ASD appear in the age between 2-3 frequently and these signs are prevalent among boys rather than girls, four times as large for boys than girls [2]. Interviews and behavioral observations are formal methods to diagnose ASD [3]. Due to the complexity of ASD symptoms and different subtypes of ASD, these methods last longer than expected. Brain imaging can be a suitable alternative for these conventional methods [4].

Rest state functional magnetic resonance imaging (R-fMRI) measures the regional interactions in brain while a person does not do a special thing. These measures are blood oxygen level dependence (BOLD) measurements which help to recognize interactions among different parts of brain, and reveal how/which different brain regions relate to each other [5]. ASD patients are unable to show normal reactions in social interactions or normal emotion responses. Thus, R-fMRI can help to detect the differences between brain patterns in a person who is suffering from ASD and normal healthy person (HP). There are several studies on diagnosing of ASD based on the information driven from R-fMRI data. In [6], the authors used regional homogeneity from fMRI data as features. Important features were selected by Chi-square feature selection, and then classified by metacognitive radial basis function classifier where 70-80% accuracy rate was obtained. Whole brain functional connectivity in the Slow-5 and Slow-4 frequency bands, 0.1-0.027 Hz and 0.027-0.073 Hz, respectively, were used as features in [7]. Support vector machine (SVM) was employed as classifier and the accuracy of 79.17% was reported. The authors in [8] got measures of pairwise functional connectivity of multi-site driven data from 7266 regions of interests (ROIs). Leave-one-out classifier was used and the obtained classification accuracy was 60%. In [9] using pipelines for extracting biomarkers from R-fMRI data was proposed. The participant-specific connectomes were assumed to learn patterns of connectivity between ASDs and HPs. The obtained classification rate was 67%.

The authors in [8] and [9] used their own templates to extract ROIs. Recently, different fMRI preprocessed data were published from the autism brain imaging data exchange (ABIDE) which helped to extract ROIs based on atlases.

fMRI data are four-dimensional (4D). To reduce the dimension from 4D to two-dimensional (2D), the average of the voxels in each ROI is calculated. To find the ROIs, atlases can be used as a template. Eickhoff-Zilles (EZ) is an atlas derived from the max-propagation atlas distributed by the SPM anatomy toolbox [19]. It represents 116 ROIs from whole brain which are used in this paper to extract time courses. Feature extraction is the most important step in image processing. The proposed feature extraction step contains two parts. First, Ripplet II transform [10] is applied on data. Then, the output of Ripplet II transform is partitioned into smaller non-overlapped parts, and cumulants [11] of each part are calculated as final features. Two-sample t-test (Ttest2) is used to find discriminative features [12]. Selected features are normalized and then classified by SVM [13]. The results show 91.67% classification accuracy.

The rest of the paper is organized as follows. Section 2 describes fMRI data, structure of the methods for feature extraction and feature selection, and classification method. In Section 3, the experimental results and the comparison of the proposed method with previous works are provided. Section 4 is a summary and conclusion of the paper.

2. Materials and Proposed Method

Fig. 1 shows the overall block diagram of the proposed method. As observed, at first, R-fMRI data are preprocessed with SPM8 [20], and then ROIs are extracted based on EZ atlas. The next two steps are feature extracting steps that include Ripplet II transform and calculating higher order cumulants. Next, the effective features are selected by two-sample T-test, and then normalized to the range [0 1]. The last step is classification. In the following each part is explained in detail.

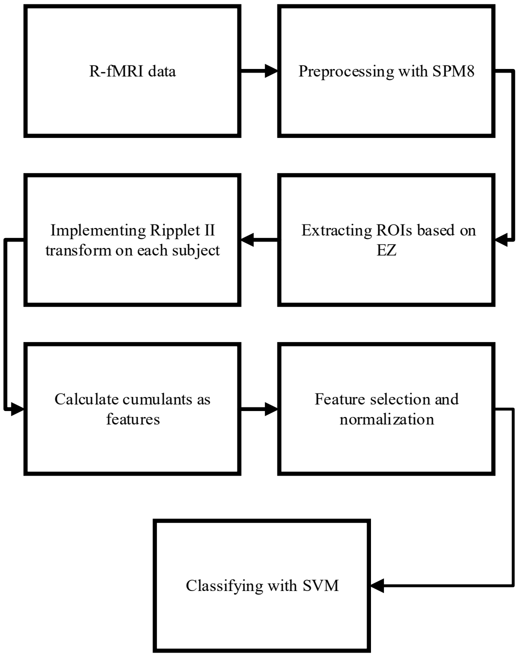


Fig. 1. Block dtiagram of the proposed method.

2.1. R-fMRI Data and Preprocessing

The R-fMRI data used in this paper are aggregated by the ABIED, Olin, Institute of living at Hartford Hospital (OLIN) [21]. Table 1 contains the details of the OLIN dataset. Due do the head motions and incomplete information in the first volumes, the first 30 volumes are ignored in this study.

Table 1. Details of the dataset used in this study.

Institute	OLIN
MRI vendor	Simens
TR (msec)	1500
TE (msec)	27
Voxel Size (mm)	3.43×3.43×4
Volumes	210
No. of ASDs	20
No. of HPs	16

Before extracting ROIs, data have to be preprocessed. The preprocessing step includes: realignment and reslicing, and normalization. All preprocessed procedure is performed with SPM8 [20].

2.2. ROI Extraction

To avoid the high dimensionality of R-fMRI data, extracting ROIs from time-series R-fMRI data can be very useful. EZ atlas [22] is used in this paper to extract 116 ROIs. Formally, the average of each voxel signal in each region is considered as time-series data.

2.3. Feature Extraction

Here we explain the proposed feature extraction method in detail.

2.3.1. Ripplet II Transform

If $X^i = \{x_1, x_2, ..., x_T\}$ is the fMRI time-series data for one ROI, then the matrix $\mathbf{F}$ can be defined as $\mathbf{F} = \{X^i;\ 1 \le \mathrm{i} \le 116\}_{116 \times T}$ for all ROIs. Thus, the data of each subject can be considered as an image in which each column contains the variations of each ROI during scanning. The fourier transform (FT) and wavelet transform (WT) are commonly used in image processing. Boundaries cause sigularities in image tensity. The disadvantage of FT is that one-dimensional (1D) singularities destroy the sparsity of FT series [14]. While WT overcomes FT constraints, it can not resolve 2D singularities along haphazard shaped curves [15]. To overcome the FT and WT restrictions, Ripplet II transform was introduced [10] which is defined based on generalized Radon transform (GRT). Ripplet II transform contains two main steps. First, GRT [23] converts singularities along curves into point singularities in generalized Radon (GR) domain. Then, WT is used to resolve point singularities in GR. To implement the Ripplet II transform, the matrix $\mathbf{F}$ is converted to $\mathbf{F}(\rho, \phi)$ in polar coordinates and the transform is defined as follows:

$$\begin{aligned}\mathbf{R}_{\mathbf{F}}(a,b,d,\theta) = 2\sum_{n=-\infty}^{\infty} \int a^{-0.5}\varphi\big((a-b)/a\big)\int_r^{\infty}\int f(\rho,\phi)e^{-jn\phi}d\phi \\ \times\Big(1-(r/\rho)^{2/d}\Big)^{-0.5} \times T_{nd}\left(r/\rho\right)^{1/d} d\rho e^{-jn\theta} dr\end{aligned} \quad (1)$$

where $a > 0$ denotes scale, $b \in \mathbb{R}$ represents translation, $d \in \mathbb{N}$ indicates degree, $\theta \in [0, 2\pi)$ indicates orientation and $(.)^*$ is the conjugate operation. $T_{nd}(.)$ is a Chebyshev polynomial of degree *nd*. More details are found in [10].

2.3.2. Higher Order Cumulants

Moments are features that have been frequently used. In this paper cumulants that are obtained from moments are considerded [11]. The matrix $\mathbf{R}$ is partioned into k non-overlapping parts. The second and forth order cumulants are extracted from each part. The second order cumulants $\{C_2 = [c_{2l}];\ l = 0,1\}$ and the forth order cumulants $\{C_4 = [c_{4j}];\ j = 0, 1, 2\}$ are calculated as follows [11].

$$c_{20}^{k} = E\left[\left(r^k\right)^2\right] - E^2\left[r^k\right] \quad (2)$$

$$c_{21}^{k} = E\left[\left(r^{k}\right)^{2}\right] \tag{3}$$

$$c_{40}^{k} = E\left[\left(r^{k}\right)^{4}\right] - 3E^{2}\left[\left(r^{k}\right)^{2}\right] \tag{4}$$

$$c_{41}^{k} = E\left[\left(r^{k}\right)^{3} \times \left(r^{k}\right)^{*}\right] - 3E\left[\left(r^{k}\right)^{2}\right]E\left[\left(r^{k}\right)\left(r^{k}\right)^{*}\right] \tag{5}$$

$$c_{42}^{k} = E\left[\left(r^{k}\right)^{2} \times \left(\left(r^{k}\right)^{*}\right)^{2}\right] - \left|E\left[\left(r^{k}\right)^{2}\right]\right|^{2} - 2E^{2}\left[r^{k}\left(r^{k}\right)^{*}\right] \tag{6}$$

where k is the partition number, $|.|$ is the absolute sign, r is the row vector of each sub-matrix, and $E[.]$ is the expectation operator. The final feature vector is revealed as $C = [C_2,\ C_4]$.

2.4. Feature Selection and Normalization

Two-sample t-test is a statistical indicator that can be used for feature selection [12]. This test is applied on two different groups and the effectiveness of features between two groups is calculated as follows:

$$t = \frac{\mu_A - \mu_B}{\sqrt{\frac{\sigma_A^2}{n_A} - \frac{\sigma_B^2}{n_B}}} \tag{7}$$

where μ is the sample mean, σ is the sample standard deviation, and n denots the number of samples in each class. Features with lower t (t < 0.01) are selected as effective features.

The selected features are normalized before applying to SVM. Normalizing not only reduces numerical difficulties during claculations, but also prevents attributes in greater numeric ranges from dominating those in smaller numeric ranges [18]. The features C are normalized to C_{norm} in the range $[0,\ 1]$ as follows:

$$C_{norm} = \frac{C - C_{\min}}{C_{\max} - C_{\min}} \tag{8}$$

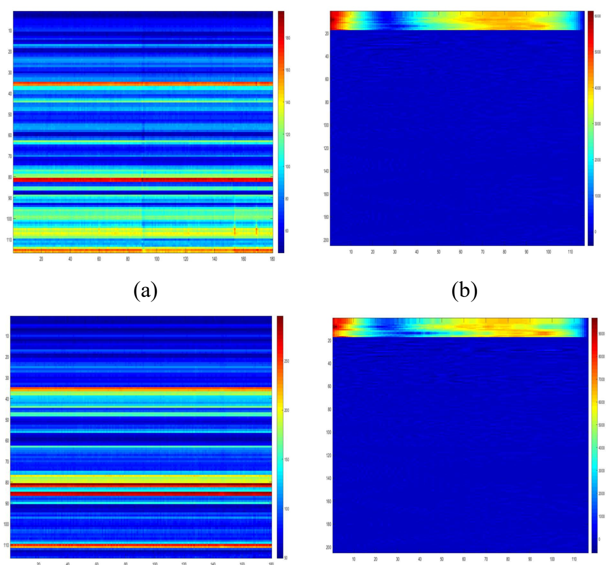
2.5. Classification

SVM is the most commonly used classifier that finds the optimal hyperplane with labeled training data which enables them to classify new examples [13]. Here, K-fold cross validation RBF SVM classifier is used which is available in MATLAB.

3. Experimental Results and Discussion

In this section, the procedure of classification of ASDs and HPs and the obtained results are presented.

By extracting time courses from R-fMRI data based on EZ atlas, each subject with 180 volumes is represented with matrix $\mathbf{F}_{116\times180}$. Then, Ripplet II transform is applied on each matrix. The Ripplet parametres are as follows: The degree (d) of the transform is set to 2 and Daubechies 4 as a wavelet function which is applied for 4 levels on each sample. Fig. 2 demonstrates the output of Ripplet II transform for two samples. The first row is a sample of ASD and the second row is a sample of HP. As shown, Fig. 2a and Fig. 2c are the results of extracting time courses based on EZ for each subject ($\mathbf{F}_{116\times180}$). Figs 2b and 2d are the Ripplet II transform coefficient matrix for those samples. The obtained matrix for each sample is partitioned into 29×41 non-overlapping sub-matrices. The second and fourth order cumulants of each part are calculated. Thus, the total number of features for each sample is $29\times41\times5$. Discriminative features are determined by two-sample T-test (t < 0.01). The selected feature vector is normalized before classification. Finally, data are classified by K-fold cross validation SVM. The Guassian kernel with 2 and 4 folds is used. The variation of classification rate for different number of features is shown in Fig. 3. The highest accuracy for 2-fold is obtained when the first 9 features with smallest t values are selected. The result is 88.89% accuracy. For 4-fold, the best result is achieved by selecting the first 10 features with smallest t values, and the method yields 91.67% classificatin accuracy.



(a) (b) (c) (d)

Fig. 2. Samples of fMRI time-series; a) a sample of fMRI time-series for ASD before Ripplet II transform, b) a sample of fMRI time-series for ASD after Ripplet II transform, c) a sample of fMRI time-series for HP before Ripplet II transform, and d) a sample of fMRI time-series for HP after Ripplet II transform.

The authors in [8] and [9] used multi-site data and different types of pipelines to figure out time courses. They achieved 60% and 67% classification accuracy, respectively. The work in [17] explored the activity of salience network among 20 ASDs and 20 HPs which were classified by independent component analysis (ICA) which resulted in 78% accuracy. In [16] the authors determined a spatial filter for projecting the covariance matrices of BOLD time-series signals. The discriminative features were extracted by a spatial feature based detection method (SFM). The method reached 77.3% classification accuracy. Table 3 compares the performance of different methods. It is observed that the proposed method significantly outperforms the previous works.

Table 3. Performance comparison of the proposed method with previous studies.

Author	Method	Accuracy (%)
Nielsen et al. [8]	Pairwise functional connectivity measurements from a lattice of 7266 ROIs obtained and grouped into multiple bins	60
Abraham et al. [9]	Using different pipelines to extract the most predictive biomarkers and finding differences of articipant-specific connectomes patterns	67
Subbaraju et al. [16]	Spatial feature besed detection method	77.3
Uddin et al. [17]	Salience network activation exploring	78
Proposed method	**Ripplet II transform and higher order cumulants as features**	2-fold SVM **88.89**
		4-fold SVM **91.67**

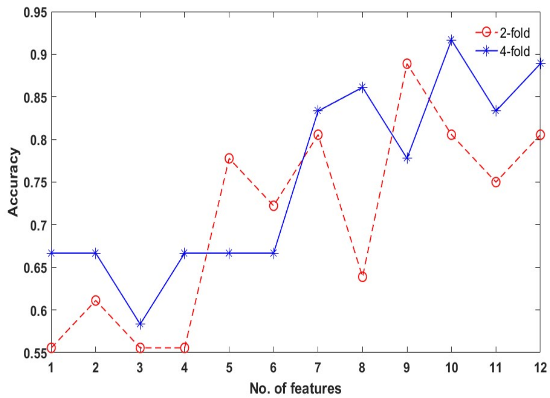


Fig. 3. Results of classification for different number of features with 2-fold SVM and 4-fold SVM.

4. Conclusion

In this work, EZ atlas was used to extract time-seies of 116 ROIs which reduces the dimension of R-fMRI data. Ripplet II transform was performed due to its robustness in exploring boundary variation in time-series data. Additionally, extracting cumulants from Ripplet II transform coefficiants improves the performance. After selecting effective features by two-sample T-test, the selected features were normalized and then were classified by SVM. With only 10 features the proposed method achives 91.67% classification accuracy which shows significant improvement in comparison with previous studies.

References

[1] D. R. Simmons, A. E. Robertson, L. S. McKay, E. Toal, P. McAleer, and F. E. Pollick, "Vision in autism spectrum disorders," *Vision research,* vol. 49, no. 22, pp. 2705-2739, 2009.

[2] M. D. Kogan *et al.*, "Prevalence of parent-reported diagnosis of autism spectrum disorder among children in the US, 2007," *Pediatrics,* vol. 124, no. 5, pp. 1395-1403, 2009.

[3] C. Lord *et al.*, "The Autism Diagnostic Observation Schedule—Generic: A standard measure of social and communication deficits associated with the spectrum of autism," *Journal of Autism and Developmental Disorders,* vol. 30, no. 3, pp. 205-223, 2000.

[4] M. A. Just, V. L. Cherkassky, A. Buchweitz, T. A. Keller, and T. M. Mitchell, "Identifying autism from neural representations of social interactions: neurocognitive markers of autism," *PloS one,* vol. 9, no. 12, p. e113879, 2014.

[5] C. Rosazza and L. Minati, "Resting-state brain networks: literature review and clinical applications," *Neurological Sciences,* vol. 32, no. 5, pp. 773-785, 2011.

[6] S. Vigneshwaran, B. Mahanand, S. Suresh, and N. Sundararajan, "Using regional homogeneity from functional MRI for diagnosis of ASD among males," in *Neural Networks (IJCNN), 2015 International Joint Conference on*, 2015, pp. 1-8: IEEE.

[7] H. Chen *et al.*, "Multivariate classification of autism spectrum disorder using frequency-specific resting-state functional connectivity—a multi-center study," *Progress in Neuro-Psychopharmacology and Biological Psychiatry,* vol. 64, pp. 1-9, 2016.

[8] J. A. Nielsen *et al.*, "Multisite functional connectivity MRI classification of autism: ABIDE results," *Frontiers in Human Neuroscience,* vol. 7, 2013.

[9] A. Abraham *et al.*, "Deriving reproducible biomarkers from multi-site resting-state data: An Autism-based example," *NeuroImage,* vol. 147, pp. 736-745, 2017.

[10] J. Xu and D. Wu, "Ripplet-II transform for feature extraction," in *Proc. SPIE Visual Communications and Image Processing (VCIP)*, 2010.

[11] M. Mirarab and M. Sobhani, "Robust modulation classification for PSK/QAM/ASK using higher-order cumulants," in *Information, Communications & Signal Processing, 2007 6th International Conference on*, 2007, pp. 1-4: IEEE.

[12] Y. Saeys, I. Inza, and P. Larrañaga, "A review of feature selection techniques in bioinformatics," *Bioinformatics,* vol. 23, no. 19, pp. 2507-2517, 2007.

[13] V. Kecman, *Learning and soft computing: support vector machines, neural networks, and fuzzy logic models*. MIT press, 2001.

[14] J. Xu, "Sparsity Based Image and Video Processing," University of Florida, 2010.

[15] S. Mallat, *A wavelet tour of signal processing: the sparse way*. Academic press, 2008.

[16] V. Subbaraju, M. B. Suresh, S. Sundaram, and S. Narasimhan, "Identifying differences in brain activities and an accurate detection of autism spectrum disorder using resting state functional-magnetic resonance imaging: A spatial filtering approach," *Medical image analysis,* vol. 35, pp. 375-389, 2017.

[17] L. Q. Uddin *et al.*, "Salience network–based classification and prediction of symptom severity in children with autism," *JAMA Psychiatry,* vol. 70, no. 8, pp. 869-879, 2013.

[18] C.-W. Hsu, C.-C. Chang, and C.-J. Lin, "A practical guide to support vector classification," 2003.

[19] Forschungszentrum Jülich, (http://www.fz-juelich.de/inm/inm-1/DE/Forschung/_docs/SPMAnatomyToolbox/SPMAnatomyToolbox_node.html).

[20] SPM, (*http://www.fil.ion.ucl.ac.uk/spm/software/spm8/*).

[21] ABIDE, *http:// fcon_1000.projects.nitrc.org/indi/abide/abideI.html*

[22] ABIDE Preprocessed, (*http://preprocessed-connectomes-project.org/abide/Pipelines.html*).

[23] K. Denecker, J. Van Overloop, and F. Sommen, "The general quadratic Radon transform," *Inverse problems,* vol. 14, no. 3, p. 615, 1998.